AN APPLICATION OF OPTIMAL TAX THEORY
TO THE REGULATION OF A DUOPOLY

by

WILLIAM JACK

No. 168    DECEMBER 1991

ABSTRACT

This paper addresses the normative issue of how tax rates should be set to control a Cournot duopoly under asymmetric information. If the marginal social value of one firm's action depends on the action of the other, taxes should be based upon the actions of both. The qualitative efficiency results of monopoly regulation carry over in a conditional Cournot sense. When pre-tax profits are independent, it is possible to implement the optimal action choices in a decentralised regulatory system, though aggregate welfare is lower. This result is used to solve the general duopoly problem when the information parameter is continuous.

Acknowledgements: This research is adapted from my D.Phil. thesis undertaken at St. John's College, Oxford. I would like to especially thank my supervisor Jim Mirrlees, as well as my examiners Paul Klemperer and David Ulpn for their help and comments. Also, thanks to Terrance Gorman and members of the Wuffield workshop, Oxford, and the Public Economics Department of the World Bank, to which earlier versions of the paper were presented, and Elizabeth Savage, Don Wright and Steffen Zies for comments. Financial support from the Rhodes Trust is gratefully acknowledged. Any errors are the author's.

CONTENTS

1. Introduction 1
2. Two Examples 3
3. The General Duopoly Problem 4
4. Discrete Information Parameter 6
5. Decentralisation 11
6. The Continuous Case with Independent Pre-tax Profits 14
7. Conclusions 16

Addendum 17
1. Introduction.

A number of recent papers (e.g. Katz and Rosen (1985), Levin (1985), Sten (1983)) have explored, under conditions of full information, how the usual results of positive tax theory, such as those of incidence analysis, are modified in the presence of imperfect competition. This paper is an attempt to address the normative issue of what optimal tax rates should be calculated in the control of oligopolistic industries, when the government is not well informed.

With full information, optimal taxes will depend on the social value of firms' actions. This may be measured in terms of some proportion of profits, and the effects of the actions on consumers. However, the former may not be observable to the government. Examples of the latter include consumer surplus in the case where the firms produce a homogeneous good, and environmental damage in the case of polluting firms. Since these welfare effects are often non-additively separable functions of the actions of the firms, the optimal tax on each firm should be based on all observed actions.

In a related literature (e.g. Holmström (1982), Nalebuff and Stiglitz (1983), Sildifer (1983)) the use of relative performance schemes has been suggested as a means of inducing competition and hence improving the welfare properties of regulatory regimes. When the relevant characteristics of firms (e.g. costs) are correlated, such schemes make full use of all information available, and generally represent an improvement. However, even when firms' characteristics are independent, multi-argument taxes may be desirable.

For example, consider a duopoly with downward-sloping demand. Consumer surplus is then a concave function of total industry output, and the marginal benefit to consumers of increasing one firm's output will depend on the output of the other firm, so that any regulatory tax should be a function of both outputs. Similarly, in a polluting duopoly, if environmental damage is convex, as is often assumed, the marginal damage of emissions from one firm depends on the emission level of the other. If taxes are designed to impose on each firm the marginal social cost of its actions, these taxes need to be sensitive to the emissions of the other firm.

Even when information is incomplete, the government can implement a Pareto efficient resource allocation as long as it is indifferent to the distribution of income between firms and consumers. This is the Groves-Loeb (1975) result applied to the oligopoly case. However, when the government faces financing constraints, taxes play the dual roles of allocating resources and raising revenue. This leads to the possibility of inefficient choices of actions by some firms. In a similar fashion to standard hidden information models, firms of the highest efficiency choose socially preferred actions, conditional on the other firm's action, but receive information rents. Conversely, the least efficient firms produce too little output/pollution, given the choice of the other firm, but are held to their reservation profits.

It is interesting to note that given the Cournot behaviour of the firms, the standard monopoly results (e.g. those of Baron and Myerson (1982)) are obtained in a Cournot sense. That is, the efficiency or otherwise of a firm's action, given the action of the other firm, follows the pattern of the monopoly case. This suggests that the monopoly results may be applied in a similar fashion to more general forms of oligopoly behaviour.

It may also be suggested that if the firms behave in a Cournot fashion, the optimal tax policy could be decentralised to a pair of Cournot type regulators, each levying taxes on its own firm so as to maximise welfare, taking the action of the other as given. However, this intuition does not hold, even when there are no externalities between firms, since the incentive compatibility constraints are modified. While the decentralised regulatory scheme we consider induces firms to choose the same actions as in the centralized case, tax revenues, and hence welfare are strictly lower.

Finally, while the decentralised regulatory regime yields lower welfare than the centralized case, it can nevertheless be used to solve the centralised problem when the information parameter is continuously distributed, and pre-tax profits are independent. This makes the results more empirically relevant.

In the next section two examples are presented. The general problem, exhibiting strategic interaction and non-separable welfare is defined in section 3, and section 4 gives its solution when the information parameter is drawn from a binary domain. Section 5 investigates the possibility of decentralising regulatory responsibility. It is shown in section 6 that the decentralisation result can be used to solve the optimisation problem when the information parameter takes on a continuum of values, as long as there is no interdependence of firms' pre-tax profits. Section 7 concludes.
2. Two examples.

(i) Homogeneous product duopoly. Consider a pair of firms, with cost functions \( c(y_i, \Phi_i) \), \( i=1,2 \), exhibiting decreasing returns to scale, where \( y_i \) is the output of firm \( i \), and \( \Phi_i \) measures its efficiency. The pre-tax benefit of firm \( i \) is given by

\[
B(y_i, y_j, \Phi_i) = P(\Phi_i; y_i, y_j) - c(y_i, \Phi_i)
\]

where \( P(\cdot) \) is the inverse demand function. If firm \( i \) pays tax \( t_i \), then post-tax profit is

\[
\pi_i = B(y_i, y_j, \Phi_i) - t_i
\]

which is constrained to be non-negative. Ignoring income effects, consumer surplus is assumed to be a concave function of aggregate output, \( S(y_1, y_2) \). If the social value of a unit of profits, \( \alpha \), is less than one, then the expected value of welfare is

\[
W = E[S(y_1, y_2) + \pi_1 + \pi_2 + \alpha(\pi_1 + \pi_2)]
\]

\[
= E[S(y_1, y_2) + B(y_1, y_2, \Phi_1) + B(y_1, y_2, \Phi_2) - (1-\alpha)(\pi_1 + \pi_2)].
\]

In general the social value of profits is unity only under conditions of perfect competition and no other taxes. See, e.g., Roberts and Spence (1977). Alternatively, \( \alpha < 1 \) if distortionary taxation is used to generate public revenue; in this case the value of a dollar of profits (on the hands of private agents) is less than the value of a dollar of government revenue. Berkow and Myerson suggest a positive interpretation of the case \( \alpha < 1 \), whereby the well-being of firm owners is not fully included in social welfare, since they live outside the jurisdiction of the regulator.

Assuming an interior solution and full information about \( \Phi_i \), the first-best optimum is characterised by Pareto optimality: outputs of each firm are such that marginal costs are equal to marginal consumer surplus. This is just the usual price equals marginal cost rule. These conditions imply that the output of each firm is a function of both \( \Phi_i \) and \( y_i \). For example, if firm 1 has low costs, it will be required to produce most of the output, as long as firm 2’s costs are not low also. Net profits at the optimum are zero, since they impact negatively upon welfare. The tax on each firm is therefore determined by this zero profit condition.

When the \( \Phi_i \) are not known the optimal outputs cannot be calculated by the government and assigned to the firm, so taxes must be used both to transfer revenue to the government and to ensure that all firms produce the optimal outputs. To this end, the taxes need to be based on the outputs of both firms.

(ii) A pair of polluting firms. Two firms, labelled \( i=1,2 \), produce emissions \( y_i \), but operate in distinct product markets. There are thus no strategic interactions in the absence of multi-argument taxes, and emissions \( y_i \) yield pre-tax benefits to firm \( i \) of \( B(y_i, \Phi_i) \), where \( \Phi_i \) measures the efficiency of the production process in terms of emissions. The function \( B \) is in fact the maximum value of gross profit given emissions \( y_i \). Inputs and output are chosen to maximise profits, and a lower level of emissions will in general lead to a reduction in the maximising level of output, as well as a change in the optimal input mix. The cost to the government of cleaning up the environmental damage is an increasing, convex function of total emissions, \( D(y_1, y_2) \). It is to finance this through taxes \( t_i \) on each firm, as well as a proportional profits tax, at a fixed rate \( \alpha \). Thus, the government’s expected revenue short-fall is

\[
-W = E \left[ \sum_{i=1}^{2} t_i (y_i, \Phi_i) \right] - \alpha E \left[ D(y_1, y_2) - \sum_{i=1}^{2} B(y_i, \Phi_i) - (1-\alpha)\pi_i \right]
\]

which is to be minimised. This has a similar form to the welfare function in the first example. Again, under conditions of full information, taxes would be based on the parameters \( \Phi_i \) and \( \Phi_2 \), to ensure zero profits and Pareto optimality. When the \( \Phi_i \) are not observable, optimal taxes should be based on the emissions of both firms, since \( D(\cdot) \) is convex. This may be interpreted within the familiar theory of externalities: each firm should be made to pay a tax at a rate equal to the marginal damage it causes. When damages are convex in total emissions, the marginal damage caused by one firm will depend on the emission level of the other, as well as its own.

3. The General Duopoly Problem.

We now consider the problem of regulating a duopoly when lump-sum taxes are unavailable. There are two firms, labelled \( i=1,2 \), which choose actions \( y_i \) in a Cournot fashion to maximise post-tax profits. Firm \( i \)'s pre-tax profits are given by the benefit function

\[\text{\footnote{This divergence is sometimes described in terms of the "cost of public funds".}}\]
\[ B(y_i, x_i, \phi_i) \]  

(1)

where \( \phi_i \) measures the efficiency (or type) of firm \( i \). The government cannot observe \( \phi_i \) (\( i=1,2 \)), but assumes they are independently and identically distributed over a domain \( \Phi_i \) with density function \( f_i(\cdot) \). Over the appropriate range of actions, the benefit function satisfies the following properties:

\[ B_i > 0 \quad B_{11} < 0 \quad B_{12} = 0 \quad B_{22} > 0 \quad B_{13} = 0 \quad B_{33} = 0 \]  

(2)

where subscripts denote partial derivatives with respect to each argument. Thus higher efficiency leads to both higher absolute and higher marginal pre-tax profits. The benefit of firm \( i \) is negatively affected by the action of firm \( j \) (e.g., through imperfectly competitive behaviour or production externalities), but this is independent of its efficiency level \( \phi_i \).

The government can levy taxes on each firm, being functions of the actions, but not the types, of both. That is, \( t_i = t(y_i, y_j) \). We assume that due to some notion of horizontal equity the same tax function is used for both firms. Net profits are \( x_i \), and the expected value of social welfare is given by

\[ W = E \left[ \sum_{i,j} B(y_i, y_j, \phi_i) - D(y_i + y_j) - (1-\alpha) \sum_i n_i \right] \]  

(3)

where \( D(.) \) is a convex function. In the consumer surplus case, \( D(.) \) is negative (so that \(-D(.)\) is positive and concave), while in the emission control example, \( D(.) \) is positive and increasing. With this second example in mind, we will refer to \( D(.) \) as the social damage function. When \( \alpha = 1 \), we have a similar situation to that of Groves and Lourie (1975), where the distribution of income between profits and consumers is important so that taxes can be used solely as incentive instruments, and the first-best is obtainable, for example with a tax \( t_i = D(y_i, y_j) \). The last term is included to ensure positive profits. For other revelation mechanisms which attain the full optimum with \( \alpha = 1 \), despite asymmetric information, see Jack, 1990, Chapter 4.) If however, the shadow price of profits is less than unity, and distribution matters, then taxes are used as revenue sources as well as incentive instruments. These two functions are not usually compatible, so the problem takes on a second-best nature. The government’s problem thus is to

\[ \max_{\phi_i, y_i} W \]  

s.t. \( B(y_i, y_j, \phi_i) - f_i(y_i) \geq 0 \quad \forall i,j \),  

and \( y_i \max B(y_i, y_j, \phi_i) - f_i(y_i) \)  

(4)

The first constraint (individual rationality) ensures that in equilibrium each firm is willing to participate in the mechanism. The incentive compatibility constraints say that the values of the action variables used in the calculation of welfare must constitute a Nash equilibrium between the firms, given the tax schedule. For this to be a realistic equilibrium concept, it is necessary to assume that the firms know each other’s type, as well as their own. If the firms do not know each other’s type, the appropriate equilibrium concept is one of Bayesian-Nash equilibrium (see, e.g. Myerson, 1980). In this case, Doms and Sappington (1985) have shown that the first-best is attainable when firms are risk neutral.

4. Discrete Information Parameter.

The simplest case is that where each firm has either a high or low efficiency level, denoted \( \phi_h \) and \( \phi_l \), respectively. For notational convenience, and without loss of generality, we take \( \alpha = 0 \).

If we assume that \( f(\phi_i) = \lambda \), then the government’s problem is to

\[ \max_{\phi_i, y_i} E[ f(y_i, y_j) ] \]  

s.t. \( B(y_i, y_j, \phi_i) - f(y_i) \geq 0 \quad \forall i,j \),  

and \( y_i \max B(y_i, y_j, \phi_i) - f(y_i) \)  

(5)

Since \( f(.) \) has binary support, the government must consider only four alternative states of nature, described by \( (\phi_h, \phi_l) \) for \( i,j=1,2 \). In general then, it will wish to implement four different combinations of actions, \( (y_{i1}, y_{j2}) \), where \( y_{ij} \) is the action level of a firm of type \( \phi_i \), when the other is of type \( \phi_j \). Thus, each firm is given a choice of four actions, and in general, an unrestricted joint choice will result in one of sixteen pairs of the form \( (y_{i1}, y_{j2}) \). We require that in
sixteen pairs of the form \((y_k, y_l)\). We require that in equilibrium, \(k = \ell\) and \(l = l\). Taxes must be such that if either or both these equalities is violated, at least one firm will be induced to change its action.\(^2\)

It is convenient to think in terms of equilibrium and disequilibrium taxes. Write

\[ t_k = B(y_k, y_k^*) \]

Then it is these equilibrium taxes, coupled with the actions \(y_k^*\) that must satisfy the usual individual rationality and incentive compatibility conditions. The full set of individual rationality constraints is

\[
\begin{align*}
\text{For a } \phi_k \text{ firm:} & \quad B(y_{kL}, y_{kL}^*) - t_{kL} \geq 0 \quad \ast \\
& \quad B(y_{kH}, y_{kH}^*) - t_{kH} \geq 0 \quad \ast \\
\text{For a } \phi_H \text{ firm:} & \quad B(y_{Hk}, y_{Hk}^*) - t_{HL} \geq 0 \\
& \quad B(y_{HH}, y_{HH}^*) - t_{HH} \geq 0 . \\
\end{align*}
\]

(6)

There are eight incentive compatibility constraints

\[
\begin{align*}
\text{For a } \phi_k \text{ firm:} & \quad B(y_{kL}, y_{kL}^*) - t_{kL} \geq B(y_{kH}, y_{kH}^*) - t_{kH} \ast \\
& \quad B(y_{kH}, y_{kH}^*) - t_{kH} \geq B(y_{kL}, y_{kL}^*) - t_{kL} \\
& \quad B(y_{HL}, y_{HL}^*) - t_{HL} \geq B(y_{Hk}, y_{Hk}^*) - t_{HL} \ast \\
& \quad B(y_{Hk}, y_{Hk}^*) - t_{HL} \geq B(y_{HL}, y_{HL}^*) - t_{HL} . \\
\text{For a } \phi_H \text{ firm:} & \quad B(y_{LH}, y_{LH}^*) - t_{LH} \geq B(y_{HL}, y_{HL}^*) - t_{HL} \ast \\
& \quad B(y_{HL}, y_{HL}^*) - t_{HL} \geq B(y_{LH}, y_{LH}^*) - t_{LH} \\
& \quad B(y_{Lk}, y_{Lk}^*) - t_{Lk} \geq B(y_{Hk}, y_{Hk}^*) - t_{Lk} \ast \\
& \quad B(y_{Hk}, y_{Hk}^*) - t_{Lk} \geq B(y_{Lk}, y_{Lk}^*) - t_{Lk} . \\
\end{align*}
\]

(7)

(8)

However, only the four starred constraints need be explicitly included. As in the monopoly case with two types (e.g., Tirole (1988)) the individual rationality constraints for \(\phi_H\) firms are

\[ (8) \]

can safely be omitted from the optimisation, and checked later (see Appendix 1 for the necessary calculations in the case of independent pre-tax profits). Finally, as regards the incentive compatibility constraints of \(\phi_k\) firms, we need only consider deviations in actions which leave the second subscript (i.e., the type of the other firm) unchanged.

Taxes enter the welfare function positively, so the individual rationality constraints on \(\phi_k\) firms imply

\[ t_{Lk} = B(y_{kL}, y_{Lk}^* \phi_k) \quad \text{and} \quad t_{Lk} = B(y_{Lk}, y_{Lk}^* \phi_k) \]

while the incentive compatibility constraints on \(\phi_H\) firms yield

\[ t_{HL} = B(y_{kL}, y_{LH}^* \phi_k) - B(y_{kH}, y_{LH}^* \phi_k) - (B(y_{Lk}, y_{LH}^* \phi_k) - B(y_{LH}, y_{LH}^* \phi_k)) \]

\[ + B(y_{Lk}, y_{Lk}^* \phi_k) - B(y_{Hk}, y_{Lk}^* \phi_k) \]

(9)

These expressions can be substituted into the objective function to give

\[ W = \lambda^2 [B(y_{Lk}, y_{Lk}^* \phi_k) - B(y_{Hk}, y_{Hk}^* \phi_k)] \\
+ 2(1 - \lambda)[B(y_{Lk}, y_{Lk}^* \phi_k) + B(y_{Hk}, y_{Hk}^* \phi_k)] \\
- B(y_{Lk}, y_{Lk}^* \phi_k) - B(y_{LH}, y_{LH}^* \phi_k) - B(y_{Hk}, y_{Hk}^* \phi_k) - D(y_{LH})] . \\
\]

(10)

First order conditions for maximisation with respect to the \(y_k\) are:

\[ \frac{\partial W}{\partial y_k} = 0 \quad \Rightarrow \quad B(y_{kL}, y_{Lk}^* \phi_k) + B(y_{kL}, y_{Lk}^* \phi_k) - D(y_{Lk}) \]

\[ = 1 - \lambda [B(y_{Lk}, y_{Lk}^* \phi_k) - B(y_{Lk}, y_{Lk}^* \phi_k)] \]

(11)

\[ \frac{\partial W}{\partial y_{kL}} = 0 \quad \Rightarrow \quad B(y_{Lk}, y_{Lk}^* \phi_k) + B(y_{Lk}, y_{Lk}^* \phi_k) - D(y_{Lk}^* \gamma_{HL}) \]

\[ = 1 - \lambda [B(y_{Lk}, y_{Lk}^* \phi_k) - B(y_{Lk}, y_{Lk}^* \phi_k)] \]

(12)

\[ \frac{\partial W}{\partial y_{LH}} = 0 \quad \Rightarrow \quad B(y_{Lk}, y_{Lk}^* \phi_k) + B(y_{Lk}, y_{Lk}^* \phi_k) - D(y_{Lk}^* \gamma_{HL}) \]

(13)

and

\[ \frac{\partial W}{\partial y_{LH}} = 0 \quad \Rightarrow \quad B(y_{Lk}, y_{Lk}^* \phi_k) + B(y_{Lk}, y_{Lk}^* \phi_k) = D(y_{Lk}^* \gamma_{HL}) \]

(14)
\[
\frac{\partial w}{\partial y_{il}} = 0 = B_i(y_{il}, y_{hl}, \phi_l) + B_i(y_{il}, y_{hl}, \phi_h) = D_i(y_{il})
\]

(1)

The second order conditions for a maximum are assumed to be satisfied. For example, if \( B_{y_{il}} < 0 \), the right hand sides of equations (12) and (13) are decreasing in the first argument. From the assumptions already made, the left hand sides are also decreasing, and we require this
decrease to be faster than that of the right hand side. This is illustrated in Figure 1, where the curve MSB(y;x) measures the (net) marginal social benefit of the action \( y \) of one firm, given that of the other. \( x \), and MIC is a measure of the marginal information cost induced by the incentive compatibility constraints.

The first order conditions have familiar properties. As in the monopoly case, the action of a high efficiency firm is socially optimal, in the sense that, given the other firm's action \( (y_{il} \) or \( y_{hl} \)), marginal damage and marginal benefit are equated (equations (14) and (15)). Similarly, equations (12) and (13) show the typical divergence of marginal benefit and cost when the firm has low efficiency, with the actions taken, \( y_{il} \) and \( y_{hl} \), below their first-best optimal levels. This parallels the results of Baron and Myerson (1982).\(^1\)

Profit levels in this duopoly case also correspond to those of the monopoly model. Firms with highest efficiency make positive after-tax profits, while the low-efficiency firms earn zero net profits. This has the familiar interpretation of an information cost to high efficiency firms.

Note that in the emission control interpretation of the model, the optimal policy results in sub-optimal emissions for \( \phi_l \) firms - that is, too much abatement. If we think of the firms as "producing abatement", then we might expect a direct transposition of the monopoly results to imply too little of this activity by low efficiency firms. The resolution of this apparent anomaly lies in the parameterization of the benefit function. In particular, low efficiency as defined in (2) does not imply high abatement costs (measured in terms of reduced benefits). Therefore, the usual results in terms of abatement levels cannot be automatically transposed. This has implications for policy, depending on whether pollution control is sought through the use of so-called end-of-pipe abatement measures, or more general changes in production plans in terms of inputs, outputs, and emissions (see Jack (1981)).

Equations (12) to (15) implicitly define the optimal actions offered by the regulator. Optimal equilibrium taxes are calculated by substituting these action levels into equations (9) and (10). In addition we must define taxes to be paid when the firms are not in an equilibrium characterized by symmetric action choices - i.e. an equilibrium of the form \( (y_i, y_j) \). To this end, let

\[
\tau(y_{iy_j}) = \tau_i \quad \text{and} \quad \tau(y_{iy_j}) = \tau_j \quad \text{for } j \neq i
\]

where

\[
\tau = \max_{i} \tau_i
\]

\(^1\) In fact, these results are generic to the optimal taxation literature. In general, it is of no use to the government to give the "best" type of agent (e.g. highest wage worker in Mirrlees (1971), lowest cost firm in Baron and Myerson (1982)) an incentive to choose less than the Pareto optimal action. Charging this agent a higher marginal tax rate will result in distortionsary losses (deadweight loss), and have no effect on revenue. However, if lower efficiency firms attract tax rates higher then the marginal damage caused by their emissions, the distortionary loss is outweighed by the increased revenue from all agents of higher efficiency.
and \( i_1 \) satisfies
\[
B(y_{i1}, y_m, \phi_{i1}) - t_{i1} > B(y_{i1}, y_m, \phi_{i2}) - t_{i2}, \quad \forall \ j \neq 1, \quad \forall \ m.
\] (17)

\( \eta \), or any higher tax, is sufficient to induce a firm to change its action if the action pair observed is not symmetric, in the sense defined above. Since welfare is a function of equilibrium actions and taxes only, the precise value of the disequilibrium tax is unimportant.

5. Decentralisation.

In this section we investigate the possibility of decentralising regulatory responsibility to a pair of regulators. We assume a specific form of decentralisation, by which each regulator ignores the policy of the other, but does consider the effect of the other firm's action on welfare through the function \( D() \). Of the two examples given above, this is more likely to occur in the environmental context, especially when firms in different industries (e.g. power generation and the production of chemical products) contribute to the same environmental problem (e.g. air quality).

Taking the action \( y' \) of firm \( j \) as fixed, and suppressing the individual rationality constraint for \( \phi_{i1} \) firms and the incentive compatibility constraint for \( \phi_{i2} \) firms, regulator 1 solves
\[
\begin{align*}
\max_{\phi_{i1}} & \left( \lambda \left( t_{i1} - D(y_{i1}, y') \right) \right) \\
\text{s.t.} & \quad B(y_{i1}, y', \phi_{i1}) - t_{i1} \geq 0 \\
& \quad B(y_{i1}, y', \phi_{i2}) - t_{i2} \geq B(y_{i2}, y', \phi_{i2}) - t_{i2}.
\end{align*}
\] (18)

In particular, each regulator does not set taxes taking the tax function of its counterpart as given, but rather taking the other firm's action as given. Thus we are not describing a Nash equilibrium between regulators with tax functions as decision variables, but more the modification by each regulator of its own firm's reaction function used when actions are chosen non-cooperatively.

The solution to the problem for each regulator is straightforward. The constraints can be substituted into the maximand to give
\[
W(y') = \lambda \left[ B(y_{i1}, y', \phi_{i1}) - D(y_{i1}, y') \right] + (1 - \lambda) \left[ B(y_{i2}, y', \phi_{i2}) - B(y_{i2}, y', \phi_{i1}) - D(y_{i2}, y') \right]
\]
where it is understood that \( y_{i1} \) and \( y_{i2} \) are functions of \( y' \). First order conditions for an optimum are then
\[
B_i(y_i, y', \phi_i) - D'(y_i + y') = \frac{1 - \lambda}{\lambda} \left[ B(y_i, y', \phi_i) - B_i(y_i, y', \phi_i) \right] - \frac{\lambda}{1 - \lambda} \left[ B(y_i, y', \phi_i) - B_i(y_i, y', \phi_i) \right]
\] (20)

and
\[
B_i(y_i, y', \phi_i) = D'(y_i + y')
\] (21)

These first order conditions can be solved to define two reaction functions, namely \( y_i(y') \) if firm \( i \) is of type \( \phi_{i1} \) and \( y_j(y') \) if it is of type \( \phi_{i2} \). Due to the symmetry of the problem, identical reaction functions can be defined for firm \( j \). By comparison of equations (20) and (21) with equations (12) to (15), we see that the equilibrium actions chosen by the firms in the decentralised mechanism (i.e. the four intersection points of the reaction functions) are different to their levels at the second-best centralised optimum.

Figure 2. The actions implemented in the decentralised regulatory mechanism coincide with those of the centralised system.

As above, tax revenue is determined by firms' benefit levels, because of the nature of the incentive constraints. Therefore, even if profits are of no social value in themselves (\( \alpha = 0 \)), higher benefits are desirable to the extent that they yield higher taxes. In the decentralised regime, if there are negative production externalities, the regulator of firm \( i \) does not take into account the effect of a higher \( y_i \) on the benefit of, and hence tax paid by, firm \( j \). From equation (20) the MSB curve in Figure 1 shifts to the right, and the action of a \( \phi_{i1} \) firm in equilibrium is higher in the decentralised case than in the centralised regime. Similarly for a \( \phi_{i2} \) firm.
On the other hand, it can be seen that when pre-tax profits are independent, the optimal second-best actions are the same in both the decentralised and centralised systems. That is,

\[ y_L(y_L) = y_L, \quad y_H(y_H) = y_H \]

and

\[ y_H(y_L) = y_H, \quad y_H(y_H) = y_H. \]

However, total welfare remains strictly less in the decentralised system due to tighter incentive constraints on high-efficiency firms. To see this, note that when both firms are of type \( \phi_H \), the tax paid by each is

\[ t_H(y_H) = \mathcal{B}(y_H, \phi_H) - \mathcal{B}(y_H, \phi_H) - \mathcal{B}(y_H, \phi_H) \]

where we have suppressed the second argument of \( \mathcal{B}(\cdot, \cdot) \). However, in the solution of the centralised problem, while the action levels associated with the \( \{y_L, y_H\} \) equilibrium correspond to \( y_L \) of the decentralised problem, the tax paid by each firm is

\[ t_H(y_H) = \mathcal{B}(y_H, \phi_H) - \mathcal{B}(y_H, \phi_H) - \mathcal{B}(y_H, \phi_H). \]

From Figure 2 it is clear that \( y_H < y_L(y_H) \) (since \( y_L = y_L(y_H), y_H > y_H \), and \( y_H(y_L) < 0 \)). Thus, \( t_H(y_H) < t_H(y_H) \). This discrepancy arises due to the fact that in the full problem, the \( \phi_H \) firm must be rewarded for not choosing the "very low" level of actions \( y_H \), when the other firm has efficiency parameter \( \phi_H \). However, in the partial problem, the \( \phi_H \) firm must be induced not to choose the relatively higher action \( y_L(y_H) \). This larger action gives correspondingly larger benefit, so the necessary financial incentive is correspondingly higher (i.e., the tax is lower). The lower tax revenue is thus a direct result of the assumption that taxes are set by each regulator, taking the other firm's actions as fixed, thereby strengthening the incentive compatibility constraints.

When the firms are of different types, the tax on the \( \phi_L \) firm is again lower in the centralised problem than in the decentralised problem. This divergence is explained in the same way as that of the \( \{\phi_H, \phi_H\} \) case above. However, the taxes on \( \phi_L \) firms are the same in both analyses since the incentive constraints for these firms do not bind. Thus, in general,

\[ t_L(y_L) = t_L(y_L), \quad t_H = t_H(y_H), \]

and

\[ t_L > t_L(y_L), \quad t_H > t_H(y_H). \]

Proposition: When firm types are independent, and \( \mathcal{V} \) has two elements, the optimal action allocations of the centralised and decentralised problems are identical, as long as pre-tax profits are independent. However, tax revenue raised in implementing these action levels, and hence welfare, is strictly higher under the centralised regime.

This result does not depend on any divergence of interests on the part of the separate regulatory bodies, since both share the same damage function. Rather, the restriction stems from the information constraints faced by each regulator. In this respect, the proposition identifies a limitation on the efficacy of co-operation between states generating common environmental problems.

6. The Continuous Case with Independent Pre-tax Profits.

The analysis of the previous section can be used to derive the solution of the duopoly problem with a continuum of types, as long as pre-tax profits are independent. Recalling that decentralised control was more likely to be found in the environmental example than the homogeneous good duopoly problem, the assumption of independent pre-tax profits can be justified by assuming the firms serve different product markets, and that there are no production externalities.

The details of the solution are relegated to Appendix 2, but it should be noted that the generalisation to the continuous case is of interest both from a theoretical and a practical perspective. It is reasonable to assume that in practice, if the efficiency parameter can take on different values (e.g., \( \phi_L \) and \( \phi_H \)), then it can take on intermediate values also. As the number of intermediate values grows, a continuum model becomes more convenient. Since any measurement of actions will be an approximation, it is hard to expect the discrete model to be empirically implemented in its pure form. It is extremely unlikely that a firm will choose any one of the four optimal actions exactly, in the binary case. Although some are performed, the explicit solution of the continuous case makes empirical estimation feasible.

The efficiency parameters of the firms are now independently and identically distributed on

---

1 Under conditions of full information, the first-best is reached with the decentralised mechanism. The tax by regulator \( i \) on firm \( j \) is \( \mathcal{B}(y_j, \phi_j) \), with \( y_j(\phi_j) \) chosen to satisfy \( \mathcal{B}(y_j, \phi_j) = D(y_j, \phi_j) \). But this is just the first order condition for the Pareto optimum in the decentralised problem with full information.
\[ \mathcal{Y} = \{ \psi, \phi \} \in \mathbb{R} \] according to the positive distribution function \( f() \). The optimisation problem is identical to that of section 3 (see (9)) with the expectation now taken over the continuous range of \( \psi \) values. Once again, we take \( \alpha = 0 \) for notational convenience, and suppress the second argument of \( B(\cdot, \cdot) \).

Before presenting the solution to this problem, we show how the usual method breaks down. The problem arises because incentive compatible action pairs cannot be easily characterised as the joint maximisers of some potential function. In orthodox cases this characterisation is possible: for instance, the output of Baran and Myerson's (1982) monopolist is that which maximises profit, while the workers in Mirrlees' (1971) income tax problem choose labour to maximise utility. Using the envelope theorem, the constraints faced by the government or principal can then be written in a form independent of the incentive instrument - i.e. independent of taxes.

The incentive constraints faced by the regulator in the continuous version of (4) now require that each action pair \((\psi, \phi) \in \mathcal{Y}(\psi, \phi, \lambda) \) constitutes a Nash equilibrium, given the tax function \( f(\cdot) \). If this is the case, the profit of firm i can be written as

\[
\pi_i(\psi, \phi) = \max_{\lambda} B(\psi, \phi) - T(\psi, \phi, \lambda) = B(\psi, \phi) - T(\psi, \phi) \cdot y_i(\psi, \phi) .
\]

(25)

The envelope theorem then gives

\[
\frac{\partial \pi_i}{\partial \psi} = B(\psi, \phi) \cdot \frac{\partial y_i}{\partial \psi} - \frac{\partial T}{\partial \psi} \cdot y_i .
\]

(26)

Hence, using the normal technique, the profit function cannot be written independent of the tax schedule, so orthodox methods are not appropriate.

The solution method relies on the decentralisation result of section 5. It is shown in Appendix 2 that that result can be generalised to any finite number of types. That is, if there are \( n \) efficiency levels, and pre-tax profits are independent, the optimal equilibrium actions under the centralised and decentralised regimes coincide. It is then shown that in the limit, the optimal actions in the continuous case can be determined indirectly by solving the decentralised problem. This is just the standard monopoly problem of Baran and Myerson (1982).

This solution identifies the second-best optimal equilibrium actions for firm i as a function of the types of both firms. That is, it characterises the function

\[
y_i(\psi, \phi) \]

(27)

The tax function \( T(\psi, \phi) \) which implements this action profile is calculated in a modified fashion to that in the binary case of section 4. (See Appendix 2.)

7. Conclusions.

The paper has explored the possibility of applying the theories of optimal taxation and regulation under asymmetric information to the control of a Cournot duopoly. When the government knows the characteristics of the firms, it should instruct them to undertake actions (outputs, emissions levels) which depend jointly on both these parameters. When such information is not available, optimal taxes should be based on the observed actions of both firms. However, without conditions of perfect competition, the shadow price of profits is less than unity, so that taxes must fund both financing (i.e. revenue raising), and incentive functions. These two roles are not completely compatible, and a trade-off exists between taxes which induce Pareto optimal actions, and higher rates to increase revenue.

When the information parameter is from a binary domain, it is possible to solve the optimisation problem directly. The qualitative results of the orthodox monopoly case then carry over in a Cournot conditional sense. When pre-tax profits are independent, the second-best optimal actions can be implemented by a particular decentralised regulatory mechanism. However, due to tighter incentive constraints, tax revenue, and hence welfare, is lower. This decentralisation result is used to solve the empirically more realistic case of a continuous information parameter, when pre-tax profits are independent.
Appendix 1.

For notational convenience, we assume that pre-tax profits are independent. We need to show that the mechanism of section 4 is incentive compatible for \( \phi_L \) firms. \( \phi_L \) firms earn zero profits, so we require

\[
0 \geq B(y_{\phi_L} \phi_L) - y_{\phi_L} = [B'(y_{\phi_L} \phi_L) - B'(y_{\phi_L} \phi_L)] - [B(y_{\phi_L} \phi_L) - B(y_{\phi_L} \phi_L)]
\]  

(A.1)

and

\[
0 \geq B(y_{\phi_H} \phi_H) - y_{\phi_H} = [B'(y_{\phi_H} \phi_H) - B'(y_{\phi_H} \phi_H)] - [B(y_{\phi_H} \phi_H) - B(y_{\phi_H} \phi_H)]
\]  

(A.2)

Now

\[
B(y \phi_H) - B(y \phi_L) = \int_{y_L}^{y_H} B'(y \phi) dy
\]

so if \( B' > 0 \), then equations (A.1) and (A.2) are satisfied if and only if \( y_{\phi_L} > y_{\phi_H} \) and \( y_{\phi_H} > y_{\phi_L} \). To verify these inequalities, first recall the first-best outputs \( y^*_L \) which satisfy

\[
B'(y^*_L \phi_L) = D'(y^*_L \phi_L)
\]

Lemma A.1: \( y_{\phi_L} > y_{\phi_H} > y_{\phi_L} > y^*_L \).

Proof: To show the first inequality, make the following definitions:

\[
\hat{y}_{\phi_L} = y_{\phi_L}
\]

and \( \hat{y}_{\phi_H} \) satisfies \( B'(\hat{y}_{\phi_H} \phi_L) = D'(y^*_L \phi_L) \).

Then

\[
B'(\hat{y}_{\phi_H} \phi_L) = B'(\hat{y}_{\phi_H} \phi_L)
\]

but

\[
\hat{y}_{\phi_H} - \hat{y}_{\phi_L} < 2y^*_L
\]

so

Bibliography.


\[ y_{in} > y_{im} \quad \text{and} \quad y_{mn} < y_{mi} \]

since otherwise

\[ B_i(y_{in} \phi_i) < B_i(y_{im} \phi_i) < D_i(y_{in} \phi_i + y_{im} \phi_i) \]

This contradicts first-best optimality, so

\[ y_{in} > y_{im} \]

The remaining inequalities are proved similarly. \(\square\)

Continuing, let us show \( y_{im} > y_{in} \). Suppose \( y_{im} \geq y_{in} \). Then

\[ B_i(y_{im} \phi_i) < B_i(y_{in} \phi_i) \]

Equations (13) and (14) then imply that

\[ B_i(y_{im} \phi_i) < B_i(y_{in} \phi_i) \quad \text{and} \quad B_i(y_{im} \phi_i) = B_i(y_{in} \phi_i) \]

since

\[ B_i(y_{im} \phi_i) < B_i(y_{in} \phi_i) \quad \text{and} \quad B_i(y_{im} \phi_i) = B_i(y_{in} \phi_i) \]

So \( y_{im} > y_{in} \) and thus

\[ y_{im} + y_{mn} > y_{in} + y_{mh} \]

\[ D_i(y_{in} \phi_i + y_{im} \phi_i) > D_i(y_{in} \phi_i + y_{mh} \phi_i) \]

But equation (14) implies that equations (A.3) and (A.4) are contradictory. Hence, \( y_{in} < y_{im} \).

But then (using equation (15))

\[ y_{in} < y_{im} \quad \text{and} \quad y_{mn} = y_{mh} \]

so \( y_{im} < y_{mn} \).

Finally, we show \( y_{im} > y_{mn} \). If \( y_{im} < y_{mn} \), then \( y_{in} < y_{mh} \) by equation (13). So

\[ y_{mn} < y_{in} < y_{mh} \]

But

\[ B_i(y_{mn} \phi_i) < B_i(y_{mh} \phi_i) \quad \text{and} \quad B_i(y_{mn} \phi_i) = B_i(y_{mh} \phi_i) \]

while

\[ D_i(y_{mh} \phi_i + y_{im} \phi_i) < D_i(y_{mh} \phi_i + y_{in} \phi_i) \]

Hence, by contradiction, \( y_{im} > y_{mn} \). By equation (12), \( y_{in} > y_{im} \), so \( y_{in} > y_{im} > y_{mn} > y_{in} \). \(\square\)

Appendix 2.

Duopoly problem with finite number of \( \phi \) values.

The solution to the duopoly problem with multiple efficiency values is completely analogous to that of the 2-\( \phi \) case. As in section 5, second-best optimal actions can be arrived at via a decentralized regulatory mechanism - again with welfare losses.

The regulator must consider \( N^2 \cdot N(1-N) \cdot (N-1) \cdot \cdots \cdot 1 = N(N+1)/2 \) different states of nature, labelled by \( (\phi_i, \phi_j) \), \( i, j = 1, 2, \ldots, N \). The optimization problem is

\[
\max_{\phi_i, \phi_j} \mathbb{E}\left[ \sum_{i, j} \left( B(\phi_i \psi_i) - D(\phi_i + \phi_j \psi_i) \right) \right]
\]

s.t. \( B(\phi_i \psi_i) > t_{ij} \quad \text{and} \quad B(\phi_j \psi_j) > t_{ij} \quad \forall \ i, j \).

(A.5)

Using the usual intuition, that individual rationality constraints bind only for \( \phi_i \) firms, and that incentive compatibility constraints bind "adjacent downwards", we can eliminate the tax terms in the maximand of equation (A.5). Thus,

\[ t_{ij} = B(\phi_i \psi_i) \quad \forall \ i, j = 1, 2, \ldots, N. \]

Then

\[ W = \sum_{i, j \neq i} \mathbb{E}\left[ \left( B(\phi_i \psi_i) - D(\phi_i + \phi_j \psi_i) \right) \right] \quad \forall \ i, j = 1, 2, \ldots, N. \]

Then

\[ W = \sum_{i, j \neq i} \left[ B(\phi_i \psi_i) - D(\phi_i + \phi_j \psi_i) \right] \]

which \( \psi_i \) is the probability firm 1 is type \( \phi_i \) and firm 2 is type \( \phi_j \), and we define \( B(\psi_i \phi_i) = 0 = B(\psi_i \phi_j) \). Hence,
where $B_j(y_{i,\kappa}, \phi_{i,\kappa}) = 0$. With a little rearrangement, the necessary first order conditions are then

$$B_j(y_{i,\kappa}, \phi_{i,\kappa}) - D_j(y_{i,\kappa}) = 0,$$

where the right hand side is zero for $i=1, N$. When the firms' efficiencies are independent, this simplifies to

$$B_j(y_{i,\kappa}, \phi_{i,\kappa}) - D_j(y_{i,\kappa}) = 0,$$

where $P_k = P_k' = 0$, and again, the right hand side is zero for $i=1, N$. Equations (A.6) are a straightforward generalisation of equations (12) to (15). The incentive compatibility constraints which we ignored above (i.e. those other than the "downward adjacent" ones) are satisfied if

$$y_{i,\kappa} = y_{i,\kappa-1} \quad \forall i = 2, \ldots, N, j = 1, \ldots, N.$$  

The same actions are implemented by a pair of decentralized taxing authorities. Suppose firm $j$'s action is fixed at $y_j$. The regulator of firm $i$ then solves

$$\max_{y_i} \sum_{j=1}^{N} p_j [L_i - D_j(y_i, y_j)],$$

s.t. $B_j(y_{i,\kappa}, \phi_{i,\kappa}) - y_i \geq 0$

and $B_j(y_{i,\kappa}, \phi_{i,\kappa}) - y_i \leq B_j(y_{i,\kappa}, \phi_{i,\kappa}) - t_i \quad \forall j = 1, \ldots, N.$

Using the constraints in the usual fashion, the maximand can be made independent of the taxes $t_i$. Differentiating with respect to $y_i$ yields first order condition which implicitly define reaction functions for each firm type, $y_i(y_j)$. The equilibria associated with these reaction functions are identical to the solutions of equations (A.6). This establishes the result.

Duopoly problem with continuum of $\phi$ values.

Following a series of preliminary definitions, we proceed in four steps. Step 1 shows that the optimal actions and taxes of the centralized duopoly problem converge, in a sense to be made explicit, to the solution of a "modified" centralized continuous problem. Step 2 shows that the solution to the decentralized duopoly problem approaches that of the decentralized continuous problem. Step 3 establishes the solution to the modified centralized continuous problem, and step 4 confirms that the solutions to the modified and original centralized continuous problems are identical.

Definitions

A partition, $S$ of $\mathcal{F}$ is a set

$$\{x_1, x_2, \ldots, x_{N+1}\} \in \mathcal{F}$$

with $x_i < x_j$ iff $i < j$, and $x_1 \approx x^-, x_{N+1} \approx x^+$.

Given a partition $S$ of $\mathcal{F}$, define the induced density function on $S(\{x_i\})$ by

$$f(s, t_{x_{i+1}}) = \frac{1}{n} N \left( \sum_{i=1}^{n} \phi(t_{x_i}) \right) \Delta t_{x_i}.$$

Given a partition $S$ of $\mathcal{F}$, let welfare restricted to $S$ be

$$W = \sum_{i=1}^{N} \left[ \sum_{j=1}^{N} f(s, t_{x_{i+1}}) - D(x_i, y_j) + y_j \cdot (y_{i,\kappa}, \phi_{i,\kappa}) \right]$$

$$= \sum_{i=1}^{N} \left[ \sum_{j=1}^{N} f(s, t_{x_{i+1}}) - D(x_i, y_j) + y_j \cdot (y_{i,\kappa}, \phi_{i,\kappa}) \right] \Delta t_{x_i} f(s, t_{x_{i+1}}).$$

As usual, welfare on $\mathcal{F}$ is

$$W = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, \phi) \cdot D(x, y_{i,\kappa}) \cdot y_{i,\kappa} \cdot (y_{i,\kappa}, \phi_{i,\kappa}) \Delta x \Delta \phi.$$

where $R(x, \phi)$ depends on $D(x)$ and $f(\phi)$. For fixed $y_i$ and a partition $S$, define the conditional welfare on $S$ by
\[ W_f(y') = \sum_{i,j} \left[ R_i(\phi_i^*) - R_i(\phi_i) \right] y_{ij}(\phi_i) \]  \hspace{1cm} (A.10)

and the conditional welfare on \( \mathcal{S} \) by

\[ W_f(y') = \int_{\mathcal{S}} [R(y(y') + \delta)](\phi) \, d\phi . \]  \hspace{1cm} (A.11)

Now, recalling the centralized ("full") and decentralized ("partial") regimes, for any partition \( S \), define the following optimization problems:

Problem \( F \)

\[ \max_{\phi, \phi'} W_f \]
\[ \text{s.t.} \quad B(y(\phi,\phi'), \phi') - R(\phi,\phi') \geq 0 \]
and \( B(y(\phi,\phi'), \phi') - R(\phi,\phi') \geq B(y(\phi,\phi'), \phi') - R(\phi,\phi') \)
\[ \forall \phi, \phi' \in S. \]

Problem \( F' \)

\[ \max_{\phi, \phi'} W_f \]
\[ \text{s.t.} \quad B(y(\phi,\phi'), \phi') - R(\phi,\phi') \geq 0 \]
and \( B(y(\phi,\phi'), \phi') - R(\phi,\phi') \geq B(y(\phi,\phi'), \phi') - R(\phi,\phi') \)
\[ \forall \phi, \phi' \in S. \]

Problem \( P \)

\[ \max_{\phi, \phi'} W_f \]
\[ \text{s.t.} \quad B(y(\phi,\phi'), \phi') - R(\phi,\phi') \geq 0 \]
and \( B(y(\phi,\phi'), \phi') - R(\phi,\phi') \geq B(y(\phi,\phi'), \phi') - R(\phi,\phi') \)
\[ \forall \phi, \phi' \in S. \]

Finally, if \( S \) is a partition of \( \mathcal{S} \), let \( y^*(\cdot, \cdot) \) denote the solution of \( F \), and \( y'^*(\cdot, \cdot) \) the solution of \( F' \).

Step 1
First extend the functions \( y^* \) and \( y'^* \) to \( \mathcal{S} \times \mathcal{S} \) by defining

\[ y^* : \mathcal{S} \times \mathcal{S} \to \mathbb{R} \]
\[ (\phi_1, \phi_2) \mapsto y^*(\phi_1, \phi_2) \]
\[ (\phi_1, \phi_2) \mapsto y'^*(\phi_1, \phi_2) \]  \hspace{1cm} (A.12)

\[ \forall \phi_1, \phi_2 \in \mathcal{S} . \]

Lemma A.2: If \( y'(s, s') \) is non-decreasing in \( s_1 \), then \( (y^*, F) \) satisfy the constraints of problem \( F' \).

Proof: Given \( (\phi_1, \phi_2) \),

\[ B(\phi_1, \phi_2) - R(\phi_1, \phi_2) - B(\phi_1, \phi_2) - R(\phi_1, \phi_2) \]
\[ \forall \phi_1 < \phi_2 . \]
The last inequality follows since \(\{y^i,t^i\} \) satisfy the individual rationality constraints of problem \( F_y \). Also,

\begin{align*}
B(y^i(t^i_s, s^i_s), t^i_f) - I^i(t^i_s, s^i_s) 
&= B(y(t^i_s, s), t^i_f) - I^i(t^i_s, s) 
&= B(y^i(t^i_s, s^i_s), t^i_f) - I^i(t^i_s, s^i_s) 
&= B(y(t^i_s, s), t^i_f) - I^i(t^i_s, s) 
&= \forall \phi_i \in \Phi_i, \exists \phi_f \in \Phi_f.
\end{align*}

The second inequality follows since \( y^i(s, s') \) is non-decreasing in \( s' \). But, the left hand side of this last expression is just

\[ I^i(\phi^i_f, \phi^i_f) - I^i(\phi^i_s, \phi^i_f) \]

Lemma A.3: \( \forall \epsilon > 0, \exists \) a partition \( S \) of \( \mathcal{S} \) such that

\[ |R(y, r') - R(S, r')| < \epsilon, \]

where \( R \) and \( r' \) extend \( S \) and \( r' \) to \( \mathcal{S} \times \mathcal{S} \), and \( y^i \) and \( r' \) solve problem \( F_y \).

Proof:

\begin{align*}
R(S, r') 
&= \int_0^1 \int_0^S [t'(s, s') + I's] \cdot D(S^i(t^i_s, s), y^{i}(s, s')) \cdot f(t^i_s, s) \cdot d(t^i_s, s) 
&= \sum_{i=1}^n \sum_{j=1}^{k_i} \int_0^1 \int_0^S [t'(s, s') + I's] \cdot D(S^i(t^i_s, s), y^{i}(s, s')) \cdot f(t^i_s, s) \cdot d(t^i_s, s) 
&= \sum_{i=1}^n \sum_{j=1}^{k_i} [t'(s, s') + I's] \cdot D(S^i(t^i_s, s), y^{i}(s, s')) \cdot f(t^i_s, s) 
&= A_{y,i}.
\end{align*}

The inequality follows from the fact that \( \{y^i, t^i\} \) solve problem \( F_y \), and any \( \{y, t\} \) satisfying the constraints of \( F^i \) (such as \( \{y^i, t^i\} \)) satisfies the constraints of \( F_y \).

Now, for any partition \( R = \{r_1, \ldots, r_m\} \) of \( \mathcal{S} \), define

\[ B_y = \sum_{s=1}^{w-1} \sum_{j=1}^{w-1} \int t'(r_j, r_j) \cdot D(y^i(r_j, r_j), y^i(r_{j-1}, r_j)) \cdot f(r_{j-1}, r_j) \cdot d(r_{j-1}, r_j) \]

By definition of the (Riemann) integral, for each \( \epsilon > 0 \), there is a partition \( R_{\epsilon} \) of \( \mathcal{S} \) such that

\[ |R(y, r') - B_{y,A_{y,i}}| < \frac{\epsilon}{2}. \]

Let

\[ m = \sup_{r \in \mathcal{S}} [t'(\phi^i_s, \phi^i_f) + t'(\phi^i_f, \phi^i_f) - D(\phi^i_f, \phi^i_f) - y^i(\phi^i_f, \phi^i_f)] \]

(This supremum exists since \( t'(\phi^i, \phi^i) \) is convex and \( D(\phi^i) \) is concave and \( D(\phi^i) \) is convex.) Then, for all \( \epsilon > 0 \), there is a partition \( Q_{\epsilon} = \{q_1, \ldots, q_m\} \) of \( \mathcal{S} \) such that

\[ \sum_{j=1}^{w-1} \sum_{j=1}^{w-1} \int [t'(q_j, q_j) - t'(q_{j-1}, q_j) - D(q_j, q_{j-1})] \cdot f(q_{j-1}, q_j) \cdot d(q_{j-1}, q_j) \leq \frac{\epsilon}{2}. \]

This follows since the double sum is separable, and by definition of the integral, again. Thus

\[ |A_{y,i} - B_{Q_{\epsilon}}, i} | < \frac{\epsilon}{2}. \]

Now let \( S_{\epsilon} = R_{\epsilon} \cup Q_{\epsilon}. \) Then \( S_{\epsilon} \) is a common refinement of \( R_{\epsilon} \) and \( Q_{\epsilon}, \) so

\[ |R(y, r') - B_{S_{\epsilon}}| < \frac{\epsilon}{2} \]

and

\[ |A_{y,i} - B_{S_{\epsilon}}| < \frac{\epsilon}{2}. \]

Hence,

\[ |R(y, r') - A_{y,i}| < \epsilon. \]

But, \( \{y, t\} \) satisfies the constraints of problem \( F^i \), so

\[ R(y, r') \geq R(S, r') \geq A_{y,i} \]

But \( \{y^i, t^i\} \) satisfy the constraints of problem \( F^i \), so

\[ |R(y, r') - R(S, r')| < \epsilon. \]
So we have established that the maximal value of welfare in problem $F'$ can be approximated arbitrarily closely by the optimal welfare level attained in problem $P_S$ for some partition $S$ of $\mathcal{S}$. If $\{y',\theta'\}$ is the unique solution to problem $F'$, then $y'=y'$ and $\theta'=\theta'$ in the sup metric (since integration is continuous with this metric). Otherwise, denote the limits of $y'$ and $\theta'$ by $y'$ and $\theta'$ respectively.

**Step 2**

The solution to the partial problem with a continuous parameter range is a simple extension of the monopoly case (see Baron and Myerson (1982), Jack (1990), Chapter 3). With $y'$ fixed, the damage function faced by the regulator of firm $i$ is $D(y',y)$. The first order conditions for $y'=y(\phi_{i},y')$ are then

$$R_i(y',\phi_{i}) - D(y',y') = \left(\frac{1-F(y',\phi_{i})}{F(y',\phi_{i})}\right) B_i(y',\phi_{i}) \tag{A.13}$$

Similarly, for $y'$ as a function of $y'$ and $\phi_{i}$. We assume that for each $y'$, the function $y(y',\phi)$ defined by (A.13) is non-decreasing in $\phi_{i}$. If not, the optimal action function is non-differentiable, and the problem must be convexified (see Baron and Myerson (1982), Guesnerie and Jaffin (1984)).

On the other hand, given $y'$, if $S=(s_1,...,s_N)$ is a partition of $S$, the first order conditions for $y'=y(\phi_{i},y')$ in problem $P_S$ are

$$R_i(y',\phi_{i}) - D(y',y') \left[ \begin{array}{c} \frac{\partial \phi_{i}}{\partial y'} \\ \frac{\partial \phi_{i}}{\partial y'} \\ \frac{\partial \phi_{i}}{\partial y'} \end{array} \right] R_i(y',s_{i+1}) - B_i(y',s_{i+1}) \tag{A.14}$$

Now if $S_m$ defined as the partition of $\mathcal{S}$ into $N$ equal units, then it is clear that as $N\to\infty$ equation (A.14) becomes equation (A.13). Thus, the solution to problem $P$ can be approximated by the solution of problem $P_S$ arbitrarily closely.

**Step 3**

The first part of this appendix argued that the actions associated with the solutions to the centralized and decentralized problems with discrete parameter ranges are identical. Within our present framework, the actions in $F$ and $F_S$ coincide. But then, the actions (but not the taxes) in the solution of problems $F'$ and $P$ must also coincide. That solution is just the simultaneous (Nash) solution to equation (A.13) and the corresponding equation for $y''$. Call the solution $y_1 = \xi(\phi_{i},y_{i})$ for $i = 1, 2, \ldots, n_j$.

Finally, we must show that the solution defined of the modified problem $F'$ coincides with the solution of problem $F$. Note that the constraints of $F'$ are stronger than those of $F$. Then we have the following.

**Lemma A1:** The function $\xi(\cdot,\cdot)$ satisfies the constraints of problem $F$. In particular, "full" incentive compatibility is satisfied.

Proofs $y'(\phi_{i},y')$ and $y''(\phi_{i},y'')$ are "fully" incentive compatible in problem $P$. Thus, $y''(\xi(\phi_{i},y_{i}))$ are "fully" incentive compatible in problem $F'$.

**Proposition A1:** The optimal actions in the control of a duopoly with independent and continuous efficiency parameters are given by $y'' = \xi(\phi_{i},y_{i})$, where $\xi(\cdot,\cdot)$ is the optimal action function derived from the simultaneous solution of the two continuous partial problems.

Proof: Lemmas A.2 to A.4.

The Optimal Tax Function.

We are now in a position to explicitly define the tax, $t(y_{1},y_{2})$ which implements the optimal action function $\xi(\cdot,\cdot)$. (Clearly $t(y_{1},y_{2})$ is defined similarly.) Firstly, define the function $\Phi_{i}$ by

$$\Phi_{i}:\mathcal{S} \to \mathbb{R}, \quad \Phi_{i}(\phi_{i},\varphi_{i}) = \xi(\phi_{i},\varphi_{i}),$$

s.t. $\Phi_{i}(\phi_{i},\varphi_{i})$ satisfies $y_{i} = \xi(\phi_{i},\varphi_{i},y_{i})$.

Now, let

$$t(\phi_{i},y_{i}) = B(\xi(\phi_{i},\phi_{i},y_{i});\phi_{i})$$

and

$$t(\phi_{i},y_{i}) = r(\phi_{i},y_{i}) = \int_{\phi_{i}}^{\phi_{i}} g(\xi(\phi_{i},\phi_{i},y_{i});\phi_{i}) \, d\phi_{i} + \phi_{i}.$$  \tag{A.17}

Next define a map $\psi_{i}$ by

$$\psi_{i}(\phi_{i},\varphi_{i}) = \xi(\phi_{i},\phi_{i},\varphi_{i}).$$
\( \psi_2 : \mathbb{R} \times \mathbb{R} \to \mathcal{F} \)

\( \psi_3(\gamma_1, \gamma_2) \) satisfies \( \gamma_1 = \xi((\phi_2, \phi_3, \gamma_2)) \)
and \( \gamma_2 = \xi((\phi_3, \gamma_2), \phi_3) \in \phi_2 \).

\( \psi_2(\gamma_1) \) picks out the \( \phi_2 \) component of the intersection of the curves \( \gamma_1 = \text{constant} \) and \( \gamma_2 = \text{constant} \) in \( \mathcal{F} \times \mathcal{F} \). Finally, let

\( r(\gamma_1, \gamma_2) = r(\psi_2(\gamma_1, \gamma_2), \gamma_2) \).

(1.18)

Proposition A.2: The optimal tax on a duopoly with a continuous parameter range, at a function of both firms' actions, is given by equation (1.18).

Proof: We proceed in three steps:

(i) \( \xi(\cdot, \cdot) \) implements \( \xi(\cdot, \cdot) \).

For fixed \( \gamma_1 \) and \( \phi_2 \), we show that

\( \gamma_2 = \xi((\phi_2, \phi_3, \gamma_2)) \) maximises \( B(\phi_2, \phi_3) - r(\gamma_1, \gamma_2) \).

The first order conditions are

\[ B_j(\gamma_1, \gamma_2) = \frac{\partial B_j(\gamma_1, \gamma_2)}{\partial \phi_2} \frac{\partial \phi_3}{\partial \phi_2} \phi_3. \]

Now

\[ \frac{\partial B_j(\gamma_1, \gamma_2)}{\partial \phi_2} \left( \frac{\partial \xi(\gamma_1, \gamma_2)}{\partial \phi_2} \right) = 1 = \frac{\partial B_j(\gamma_1, \gamma_2)}{\partial \phi_3} \frac{\partial \phi_3}{\partial \phi_2} \]

and

from the definition of \( \psi_3 \). Hence, the first order conditions are

\[ B_j(\gamma_1, \gamma_2) = B_j(\xi(\gamma_1, \gamma_2), \phi_3, (\phi_3, \gamma_2), \phi_3(\gamma_1, \gamma_2)) \]

which is satisfied by \( \gamma_1 = \phi_3 \) and \( \gamma_2 = \xi(\phi_1, \phi_3) \), as required. Concavity of the benefit function means that the first order conditions are sufficient for a maximum.

(ii) If \( r(\gamma_1, \gamma_2) > 0 \) for some \( \gamma_1, \gamma_2 \) then there is a \( \phi_2 \) such that \( \gamma_2 = \xi((\phi_2, \phi_3, \gamma_2)) \) is not chosen by firm 2. There are a number of alternatives to consider:

(a) \( r(\cdot, \cdot) = k + t(\cdot, \cdot) \) an affine shift. This will violate individual rationality for at least the lowest \( t(\cdot, \cdot) \).

(b) \( t(\cdot, \cdot) \) is differentiable at \( (\gamma_1, \gamma_2) \), \( t(0, \gamma_2) = t(0, \gamma_2) \) and

\[ \frac{\partial t}{\partial \gamma_2}(0, \gamma_2) = \frac{\partial t}{\partial \gamma_2}(0, \gamma_2). \]

Then for some \( \gamma_2 = t(0, \gamma_2) \)

\[ \frac{\partial t}{\partial \gamma_2}(0, \gamma_2) = \frac{\partial t}{\partial \gamma_2}(0, \gamma_2). \]

(using the right hand derivative if necessary, or \( t(\cdot, \cdot) \) has a jump discontinuity at \( (\gamma_1, \gamma_2) \). In either case, a firm of type \( \phi_3 \) satisfying \( \gamma_2 = \xi(\gamma_1, \gamma_2), \phi_3(\gamma_1, \gamma_2) \) will not choose \( \gamma_2, \gamma_1 \).

(c) \( r(\cdot, \cdot) \) is differentiable at \( (\gamma_2, \gamma_2) \) with

\[ \frac{\partial r}{\partial \gamma_2}(\gamma_2, \gamma_2) = \frac{\partial r}{\partial \gamma_2}(\gamma_2, \gamma_2). \]

Clearly marginal incentives are incompatible with the implementation of \( \xi(\cdot, \cdot) \).

(d) \( r(\cdot, \cdot) \) or its partial derivatives discontinuous at \( (\gamma_1, \gamma_2) \). Again, \( \xi(\cdot, \cdot) \) will not be implemented.

(iii) Thus \( \xi(\cdot, \cdot) \) is the largest tax function which implements the action function \( \xi(\cdot, \cdot) \), so \( t(\cdot, \cdot) \) is the optimal tax. \( \square \)
Working Papers in Economics

116 G. Kingston
Theoretical Foundations of Constant-Propportion Portfolio Insurance: Some Preliminary Results; November 1988

118 W.P. Hogan
Insider Information and Market Adjustment:
November 1988

119 R. Klemm
Reinterpreting a Recent Temporally Aggregated Consumption-Cap Model: December 1988

120 C. Groenewegen
Progressive Personal Income Taxation: A Historical Perspective; December 1988

122 W.P. Hogan & L. G. Sheperd
Rubinstein's Solution of the Bargaining Problem: Some Generalisations and Extensions; December 1988

123 M. P. Blund & R. Oulton
Prudential Regulator of Bank Ownership and Control; January 1989

125 D. Mills
The Reform of Australian Aviation; June 1989

126 A. E. I. Kehnn
Is Austerity Necessary? July 1989

127 S. Gill
Labour Market Flexibility - To What End? August 1989

128 J. H. R. & J. Sheen
On Optimal Damping; September 1989

129 S. K. Gharvazi
Multiple Reputations in Finitely Repeated Games; October 1989

130 J. Sheen
International Monetary and Fiscal Policy Cooperation in the Existence of Wage Inflations; October 1989

131 E. Jones
Was the Post-War Boom a Mirage? October 1989

132 S. L. Shepp & J. Sheen
A Risk Averse Price-Setting Monopsonist in a Model of International Trade; October 1989

133 F. Gill
A Target-Wage Dilemma: Some Consequences of Incomplete Information; December 1989

134 W. P. Hogan
New Banks in Australia; December, 1989

135 R. Varoufakis
Modelling Rational Conflict: The Limits of Game Theory; February 1990

136 L. Kehnn
Strategic Investment, Competition and the Independence Result; March 1990

137 D. J. Wright
International Technology Transfer with an Anonymity and Endogeneity Research and Development; April 1990

139 D. J. Wright
International Technology Transfer and Patent Royalties; April 1990

140 P. Ganguli & S. Bhat
Optimal Mix of Urban Public Services: The Case of Three Indian Cities; May 1990

141 P. D. C. Greenaway
Alfred Marshall's Principles of Economics: A Centenary Perspective from the Antipodes; June 1990

142 J. Sheen
Real Wages and the Business Cycle in Australia; June 1990

143 J. K. Karfasik
A Model of Exchange Rate Policy: Evidence for the US Dollar-Czech Koruna Rate; August 1990

144 C. J. Karfasik & D. H. Nelson
Interest Rate Linkages within the European Monetary System; A Time Series Analysis; August 1990

145 C. J. Karfasik & D. H. Nelson
Asymmetry in the European Monetary System: Evidence from Interest Rates; September 1990

146 W. P. Hogan
International Capital Adequacy Standards; October 1990

147 J. Y. W. Chan
Shrunk Ownership: The Socialisation or Privatisation of Housing? October 1990

148 G. Butler
Contracts in the Political Economy of a Nation; November 1990

149 B. Rao
Some Further Evidence on the Policy Ineffectiveness Proposition; November 1990

150 D. J. Wright
Hidden Action and Learning-By-Doing in Monopoly Regulation and Infant Industry Protection; November 1990

151 C. J. Karfasik
Testing for Long Run Money Demand Functions in Greece Using Cointegration Techniques; November 1990

152 D. C. Hutchinson & S. Nicholas
The Internationalisation of Australian Business: Technology Transfer and Australian Manufacturing in the 1980s; November 1990

153 B. Rao
A Disequilibrium Approach to the New Classical Model; December 1990

154 J. R. Towe
The Determinants of American Equity Investment in Australia; December 1990

155 E. Jones
Economists, The State and The Capitalist Dynamic; January 1991

156 I. S. Irvine & W. A. Sims
German Solar Panels and the 6-Branch Utility Tariff; February 1991

157 B. Rao
A Model of Income, Unemployment and Inflation for the U.S.A.; February 1991

158 W. P. Hogan
New Banks: Impact and Response; March 1991

159 P. D. C. Greenaway
Decentralising Tax Revenues: Recent Initiatives in Australian Federation; April 1991

160 C. J. Karfasik
Monetary Policy and the Velocity of Money in Greece: A Cointegration Approach; July 1991

161 D. J. Wright
Disaggregation, Disequilibrium and the New Classical Model; July 1991

162 Y. Varoufakis
Postmodern Threats to Game Theory; August 1991

163 Y. Varoufakis
Freedom within Reason from Axioms to Maximal Praxis; August 1991

164 D. J. Wright
Permanent vs. Temporary Infant Industry Assistance; September 1991

165 C. J. Karfasik & A. J. Phillips
Covered Interest Parity and the Efficiency of the Australian Dollar Forward Market: A Cointegration Analysis Using Daily Data; November 1991

166 H. Jack
Pollution Control versus Abatement: Implications for Taxation Under Asymmetric Information; November 1991

167 C. J. Karfasik & A. Patil
Exchange Rate Convenience and Market Efficiency; December 1991

168 W. J. Jack
An Application of Optimal Tax Theory to the Analysis of the Regulation of a Monopoly; December 1991

Copies are available upon request.

Department of Economics,
The University of Sydney,
N.S.W. 2006, Australia.
Working Papers in Economics Published Elsewhere

15. V.B. Hall, Economic Record, 56(152), March 1980.
34. Economic Record, 60(168), March 1984.
35. Economic Record, 59(166), September 1983.
36. Economic Record, 59(166), September 1983.
41. Economic Record, 59(166), September 1983.
42. Australian Economic Papers, 22(41), December 1983.
43. Economic Letters, 12, 1983.
44. Energy Economics, 8(2), April 1986.
47. Australian Economic Papers, 23(43), December 1984.
52. Economic Record, 62(178), September 1986.
112 P. Groenewegen  
   Neoclassical Economic Theory 1870 to 1938  

113 V. B. Hall,  
   T. P. Truong &  
   V. A. Nguyen  
   Energy Economics, 42(4) October 1990

114 V. B. Hall,  
   T. P. Truong & V. A. Nguyen  
   Australian Economic Review, (87) 3'89

115 P. Gill  

116 G. Kingston  

117 V. B. Hall &  
   D. R. Hills  

118 W. P. Hogan  
   Abacus, 25(2), September 1989

119 P. Groenewegen  
   Flattening the Tax Rate Scale: Alternative Scenarios and Methodologies. (eds.) J. G. Head and K. Kremer, 1-3, 3-11, 1990

120 W. P. Hogan &  
   I. G. Sharpe  
   Economic Analysis and Policy, 19(1), March 1989

121 G. Hills  

122 F. Gill  
   The Australian Quarterly, 61(4), 1989

123 S. Lahiri &  
   J. Shean  
   The Economic Journal, 100(400), 1990

124 C. J. Karfakis  

125 C. J. Karfakis  
   Journal of Money, Credit, and Banking, 22(3), 1990

126 & D. Moschos  
   Economic Papers, 10(1), March 1991